

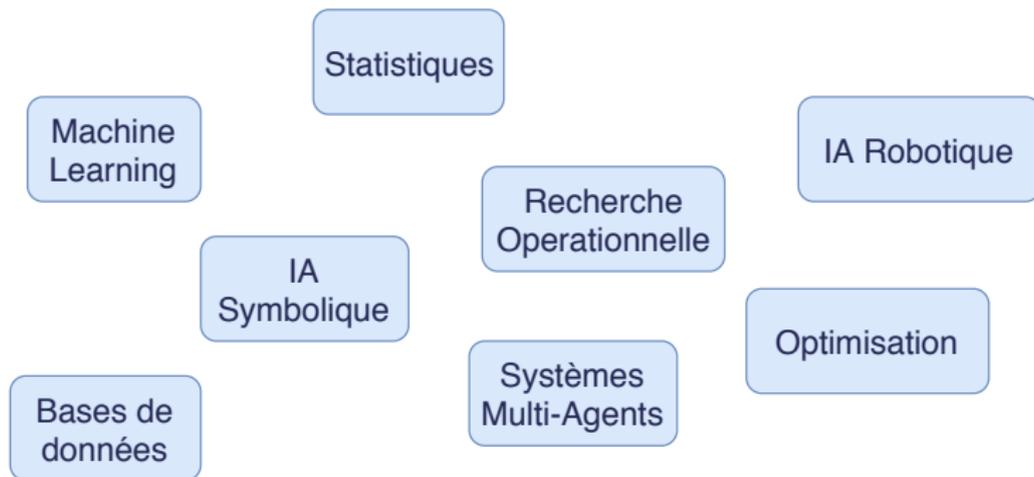
L'ENSEIGNEMENT DE L'IA À SORBONNE UNIVERSITÉ

10 Janvier 2019

Journée enseignement et formation de l'IA

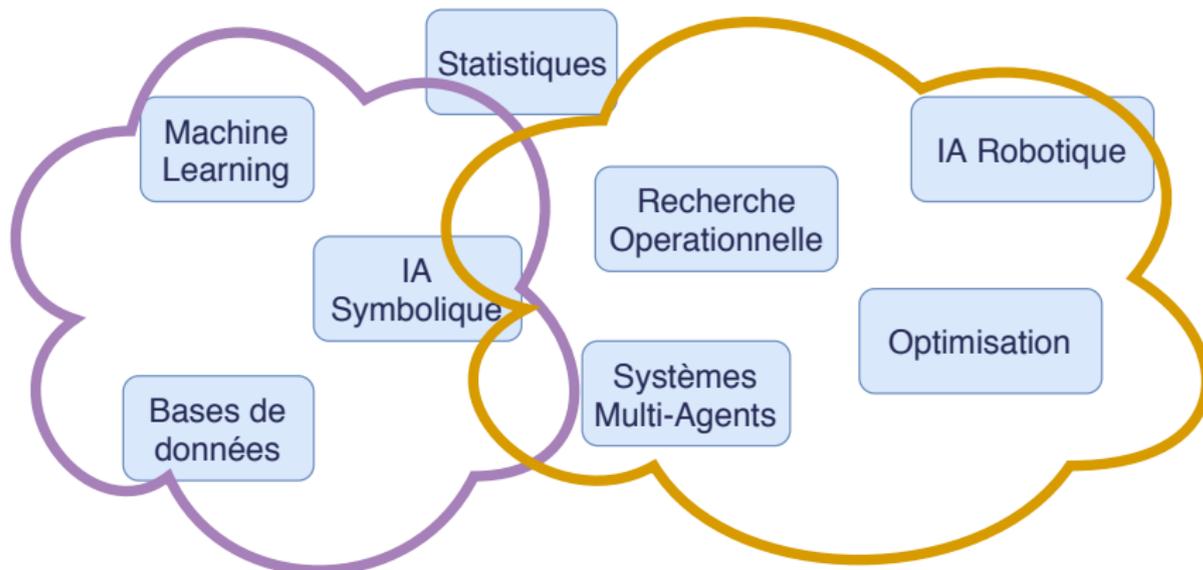
N. Bredèche, N. Maudet, B. Amann, V. Guigue

MASTER
EN SCIENCES ET TECHNOLOGIES
MENTION INFORMATIQUE



Master Informatique - Spécialité ANDROÏDE :

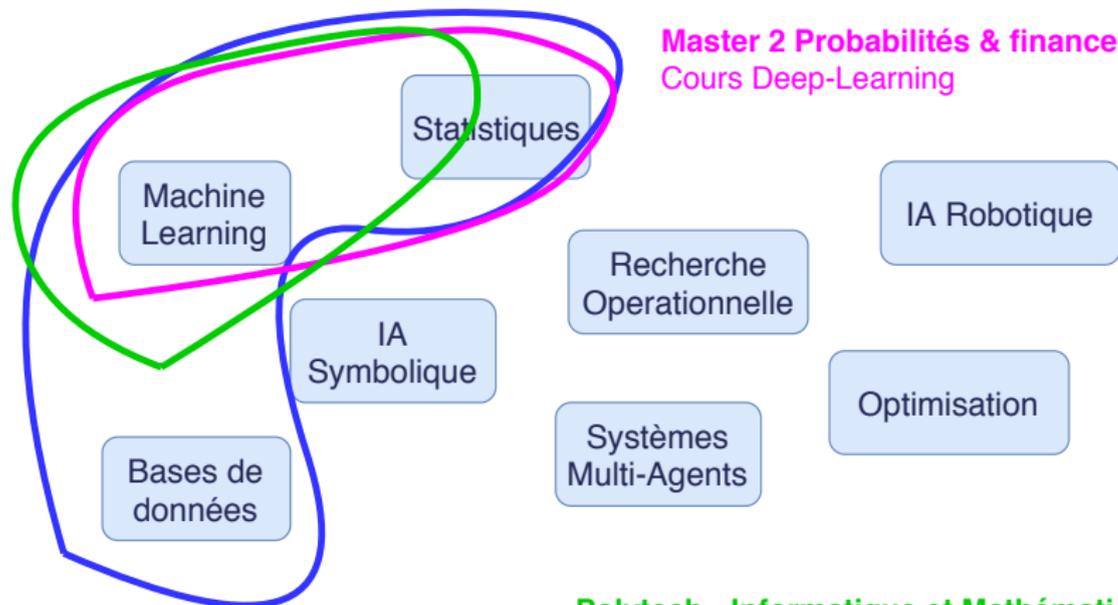
AgeNts Distribués, Robotique, Recherche
Opérationnelle, Interaction, DEcision



Master Informatique - Spécialité DAC :

Données Apprentissage Connaissances

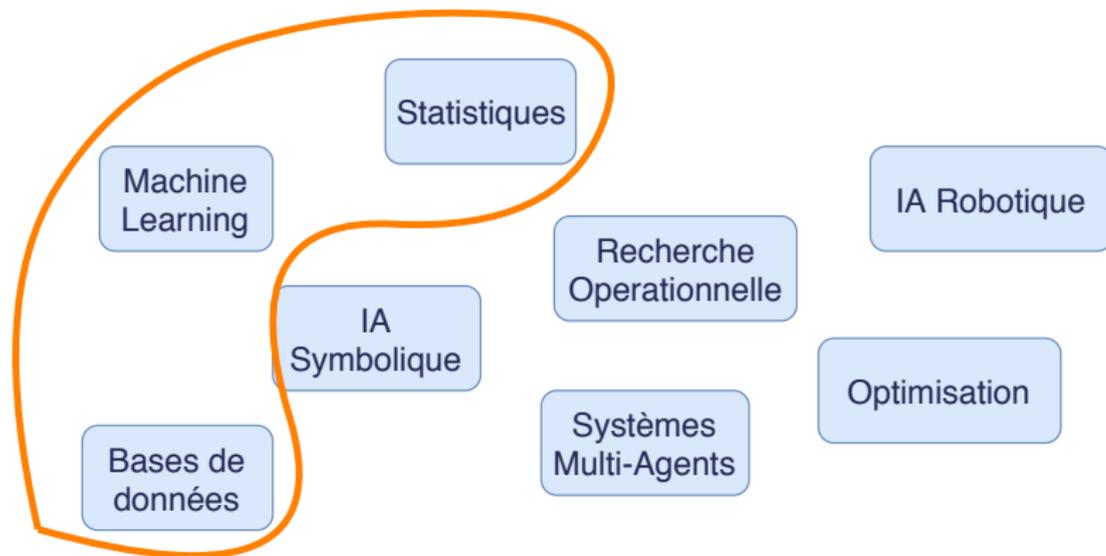
S1 en commun, S2, S3 spécifiques, S4 = stage industriel/recherche



Master 2 Probabilités & finance
Cours Deep-Learning

Master 2 Mathématiques & Applications
Mention Big-Data

Polytech - Informatique et Mathématiques Appliquées
Cours Machine-Learning



Machine Learning & Intelligence Artificielle

18 jours

Formation de 9 x 2 jours

cogérée par les départements de mathématiques et informatique

L'IA est un domaine:

- (1) principalement enseigné au niveau Master
- (2) où le doctorat est **très** valorisé

Flux d'entrée:

Idéal: Licence double cursus math-info

DAC/AND: > 60% issus de la licence d'informatique

Math/Stat: Licence de mathématiques / prépa

Des licences en mutation:

- + d'UE statistiques en info
Stats & info, Intro Data Sciences, Intro Décision/RO, Outils maths
- + d'UE d'info en mathématiques
python > Maple, Mathematica, etc...

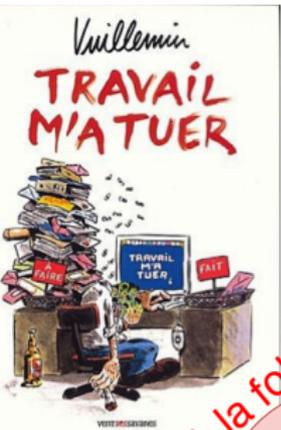
⇒ Prise de conscience des débouchés en IA...

...de la part des étudiants

Sélection en master = nouvelle donne pour les masters IA

- 20 à 30% de dossiers retenus en master (DAC/ANDROÏDE)
 - Uniquement des premiers choix sur la spécialité
 - UE L3 Stats & info réputée difficile:
faible attractivité < 2017 \Rightarrow Groupes (trop) pleins
- Fort niveau de demande du certificat Big-Data en M2 maths
- Très forte attractivité de la formation continue
(18 personnes dès l'ouverture en 2018)
- Mélange & émulation des étudiants :
SU, ENS, Telecom, Mines, ENSTA, ...

\Rightarrow Très bon niveau en sortie de M2



Un peu
Les maths:
 probas,
 stats, logique,
 ...

Compute the standard deviation for 14, 15, 19, 21, 13, 13
 1st compute mean, $\bar{x} = \frac{\sum x}{n} = \frac{14+15+19+21+13+13}{6}$
 $\bar{x} = 16$
 2nd Remember the formula for standard dev.
 $S = \sqrt{\frac{\sum (x-R)^2}{n-1}}$
 3rd Lay out table using order of operation

x	x-R	(x-R) ²
14	14-16 = -2	(-2) ² = 4
15	15-16 = -1	(-1) ² = 1
19	19-16 = 3	(3) ² = 9
15	15-16 = -1	(-1) ² = 1
18	18-16 = 2	(2) ² = 4
19	19-16 = 3	(3) ² = 9
21	21-16 = 5	(5) ² = 25
13	13-16 = -3	(-3) ² = 9
13	13-16 = -3	(-3) ² = 9
Σ	Σ(x-R) = 0	Σ(x-R)² = 59

 $S = \sqrt{\frac{\sum (x-R)^2}{n-1}} = \sqrt{\frac{59}{6}} = \sqrt{10.83} = 3.29$

A la folie
Apprendre,
Travailler,
Inventer,
réfléchir

Beaucoup
Programmer,
modéliser

```
def add5(x):
    return x+5

def dotwrite(ast):
    nodename = getNodename()
    label=symbol.sym_name.get(int(ast[0]),ast[0])
    print ' %s [%label="%s" % (nodename, label)
    if isinstance(ast[1], str):
        print '%s' % ast[1]
    else:
        print '='
    print '=';
    children = []
    for in n, children enumerate(ast[1:]):
        children.append(dotwrite(child))
    print ' %s -> (' % nodename
    for in :namechildren
        print '%s' % name,
```

Schéma type d'une UE

Cours:	Cas d'usage, Formalisation mathématique, Principaux théorèmes
TD:	Mise en équation de situations concrètes, Implémentation & réflexions algorithmiques, Démonstrations basiques
TP:	Mise en pratique jusqu'à obtention d'un système fonctionnel, formation aux bibliothèques de l'état de l'art

Evaluations: examens sur feuille, projets, bibliographie

Interventions industrielles: Facebook, Dataiku, Talend, SAP

Conserver une bonne
vision de l'état de l'art

Maintenir le **savoir
faire technique**

Entretenir les liens avec
les **industriels**

- Adossement obligatoire sur
des **équipes de recherche** fortes et larges
- Révision annuelle des supports
- Réorganisation des cours tous les 2 ans au minimum
- Risque de débauchage industriel
profil académique = profil recherché par les industriels

Les **technologies** bougent vite, les **plateformes logicielles** évoluent tous les mois, **l'industrialisation** des algorithmes de recherche se fait dans des délais incroyables (6 mois-2 ans), des **articles** reposent sur des bibliographies entières de moins de 3 ans...

- ... dans tous les secteurs de l'informatique
 - Infrastructures matérielles
 - Développement des plateformes logicielles IA
 - Exploitations des plateformes, industrialisation du traitement des données
 - Recherche algorithmique, POC
 - ... dans tous les secteurs d'activité
 - Finance, énergie, transport, robotique, IoT, industrie lourde ...
 - ... à tous les niveaux de formation L, M, D
- 75% des stages de master ont lieu en R&D
 - 1 étudiant sur 3 poursuit en doctorat (CIFRE ou académique)

⇒ Pénurie réelle de main d'œuvre aux niveaux ingénieur et docteur

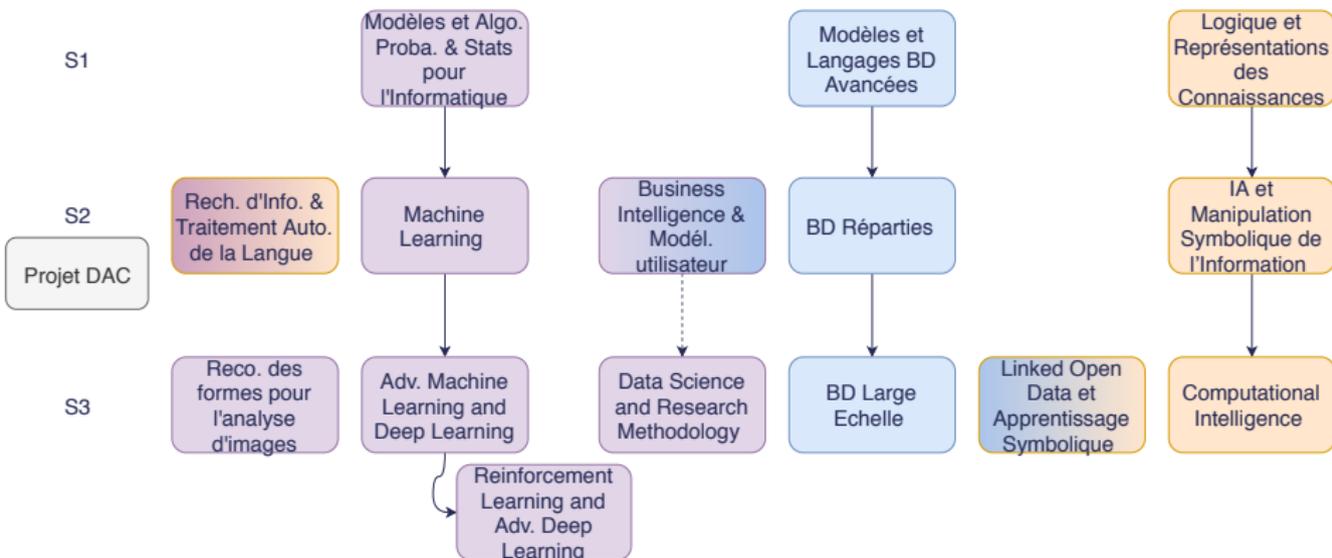
Centrée sur les données:

acquisition, traitement, stockage, valorisation des données,
raisonnement, aide à la décision

Domaines applicatifs:

- Données textuelles: classification, traduction, recherche d'informations
- Données images: reconnaissance d'objets, vision
- Séries temporelles, prédiction de flux, smart home/city/transportation, maintenance prédictive
- Traces utilisateurs, recommandation, modélisation
- Modélisation des connaissances

<http://dac.lip6.fr/master/>



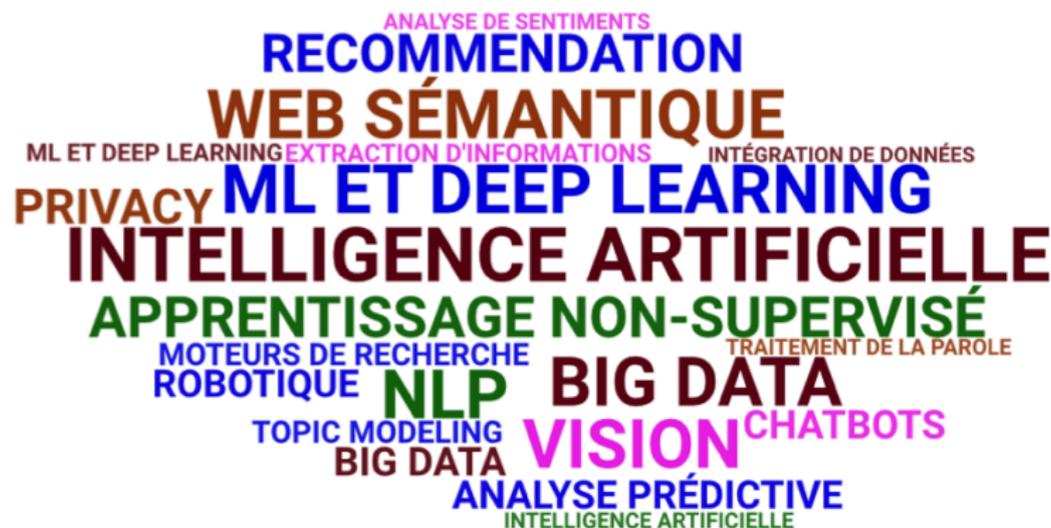
- Equilibrage Cours/TD/TP : la spécificité de DAC par rapport aux autres formations franciliennes

- ENGIE, Véolia, EDF R&D
- Crédit Agricole, Société Générale, BNP, AXA
- Cdiscount, Veesion.io, BotFul, Numen, Heuritech, Social Object, Soft.Computing, Weborama, Planorama
- ONERA, Airbus, Thales
- Orange Labs, Nokia Bell-Labs
- L'Oréal AI, SAP
- SoftBank Robotics, SYHA
- Deezer

Recherche publique:

- Institut du Patrimoine
- Hôpital Georges Pompidou, Institut de Neurosciences de la Timone, Institut du Cerveau de la Moelle épinière
- **Paris:** LIP6, LOCEAN, LTCI, CNAM, ENS **Montpellier:** LIRMM, **Grenoble:** GIPSA, LIG, **Nancy:** LORIA

Le même panorama en nuage de mots:



- A reinforcement learning-driven translation model for search-oriented conversational systems
- Classification automatique et recherche d'images par le contenu.
- Prédiction des acheteurs potentiels de véhicules sur les sites web en utilisant l'apprentissage automatique
- Classification de Pages Web
- Architecture Hautement disponible distribuée pour l'analyse de données
- Utilisation de l'ontologie pour l'intégration de donnée
- Machine learning & étude de mails
- Application du machine learning aux attaques logicielles
- Topic Modeling and Recommendation
- Système de Questions Réponses
- State Representation Learning for Reinforcement Learning
- Mise en place d'un système de rééquilibrage de stocks
- Application de méthodes de sémantique vectorielle à la classification de documents et à l'extraction de fragment
- Exploring deep vision models for general audio processing
- Détection automatique de fraude à l'assurance
- Apprentissage de comportements et aide à la décision
- Intrinsically Motivated Goal Exploration Processes
- PredictMA: A Study To Predict Early Alzheimer Disease with Resting State fMRI Data
- Prévission de l'activité journalière du site e-commerce de XXXXX afin d'améliorer les processus de l'entreprise
- Landmark Recognition in Pictures
- Validation de relations par des méthodes neuronales

MENTION INFORMATIQUE

spécialité

ANDROIDE

AGENTS DISTRIBUÉS, ROBOTIQUE, RECHERCHE
OPÉRATIONNELLE, INTERACTION, DÉCISION

Différentes filières autour de l'IA:

Environnements interactifs

Jeux vidéos, jeux sérieux, IHM, e-learning, systèmes d'information



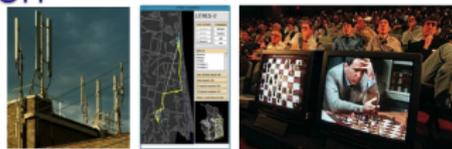
Robotique et systèmes intelligents

Systèmes multi-agents, robotique industrielle et domestique



Recherche opérationnelle et décision

Transports, énergie, logique, banques



<http://androide.lip6.fr>

M1-S1 (5 UE dont 2 de spécialité)

Modélisation, optimisation, graphes et programmation linéaire
Logique et représentation des connaissances (commune avec DAC)

10 semaines

M1-S2 (5 UE parmi 6)

UE projet (obligatoire)
Interaction homme machine
Fondements des systèmes multi-agents
Décision et jeux
Résolution de problèmes
Robotique et apprentissage (3 ECTS — UE combinée avec Anglais 3 ECTS)

10 semaines

M2-S3 (5 UE parmi 9)

<i>Ordonnancement et optim. combinatoire</i>	<i>IA pour la robotique</i>
<i>Théorie des jeux</i>	<i>Modélisation et simulation multiagents</i>
<i>Décision multi-critères ou collective</i>	<i>Environnements virtuels hautement interactifs</i>
<i>Décision dans l'incertain</i>	<i>Ingénierie des serious games</i>
<i>Coordination et consensus multiagents</i>	

12 semaines

Stage obligatoire

Académique ou industriel
France ou étranger (Japon, USA)

6 mois

Stages M2

Promo 2018 (37 étudiants M2):

- 1/3 stages académique
SU (LIP6 et ISIR), LAMSADE, LIRIS, LEME (Nanterre), AgroParisTech/INRA, NII (Japon)
- 2/3 stages industriel
ALTEN (3), Thales (2), Energisme (2), Brain Corp USA (2), DGA, Camélia Tech/B-Reputation, DGA, DXC, Eurodecision, FRAUNHOFER IPA, Interpol à Lyon, Kyriba, Manzalab, Olféo, Onera, SNCF, Sopra-Stéria, Thalès, Ubisoft



33%

Doctorat

Promo 2018:

- 9 thèses, dont:
 - 3 bourses ministérielles EDITE
 - 3 thèses CIFRE (avec le LIP6)



67%

Entreprises

Promo 2018:

- 7 embauches à la suite du stage
 - Alten, DXC, Eurodecision, Interpol, Soprasteria, Ubisoft, Braincorp (USA)
- ~100% d'embauche à 6 mois (données sur les années précédentes)

- Mécanismes de sureté de fonctionnement pour la robotique autonome
- Architecture d'agent adaptatif et autonomie à long terme
- Mécatronique et deep learning pour la récolte autonome de tomates en serre
- Détection de fraudes bancaires intra-entreprise par machine learning
- Développement d'un jeu sérieux traçable de physique
- Implémentation du bouclage macroéconomique dans un simulateur multi-agents
- Vérification formelle de logiciels embarqués à base de réseaux de neurones
- IA pour la gestion des comportements d'avatar en réalité virtuelle,
- Développement et interfaçage d'une prothèse de main
- Question Answering basés sur de l'apprentissage automatique à partir de document
- Programmation Gameplay
- Deep Learning sur un moteur de recherche visuelle
- On Randomness in Scheduling Problems
- Développement des véhicules autonomes : application au cas de la tondeuse
- Segmentation parole/musique par apprentissage profond à l'institut national de l'audiovisuel
- Re-planification temps-réel des circulations ferroviaires en cas d'aléas
- Développer de la connaissance autour des techniques d'apprentissage par renforcement dans le cadre de l'évitement d'obstacle
- Elicitation de préférences et optimisation multicritère ou multiagent en présence d'évaluations bi-polaires
- Apprentissage et Optimisation pour décisions optimales dans la gestion d'énergie
- Développement Intelligence artificielle
- Modélisation du développement du chant de Diamant Mandarin
- Longterm SLAM for mobile robots
- Approche décentralisée de régulation dans un système ferroviaire
- Hypergraphes de majorité pour le choix social computationnel
- Développement et intégration de solutions d'intelligence artificielle
- Génération de trajectoires 2D pour le suivi de zone libre par un drone
- Intelligence artificielle appliquée au pilotage des missiles
- Apprentissage par renforcement et évitement d'obstacle en robotique
- Adaptation des politiques de mobilité urbaine par l'étude de leur impact sur la qualité de l'air
- Procédures de décision incrémentales pour la décision collective sur domaine combinatoire
- Méthode robuste de construction de recommandations multicritères
- Intelligence Artificielle sur un déambulateur robotisé
- Optimisation stochastique multi-stage pour gestion optimale d'énergie
- Liage de données et transfert learning

Formations plus courtes

Tronc
commun:

Apprentissage
statistique

Outils pour le
Cloud

Big Data et ses
applications

Options selon
les
spécialités:

HPC, algo.
parallèles,
stabilité
numérique

Méthodes et
algorithmes pour
le calcul
parallèle

Massive parallel
programming on
GPU devices for
Big Data

Statistique et
apprentissage

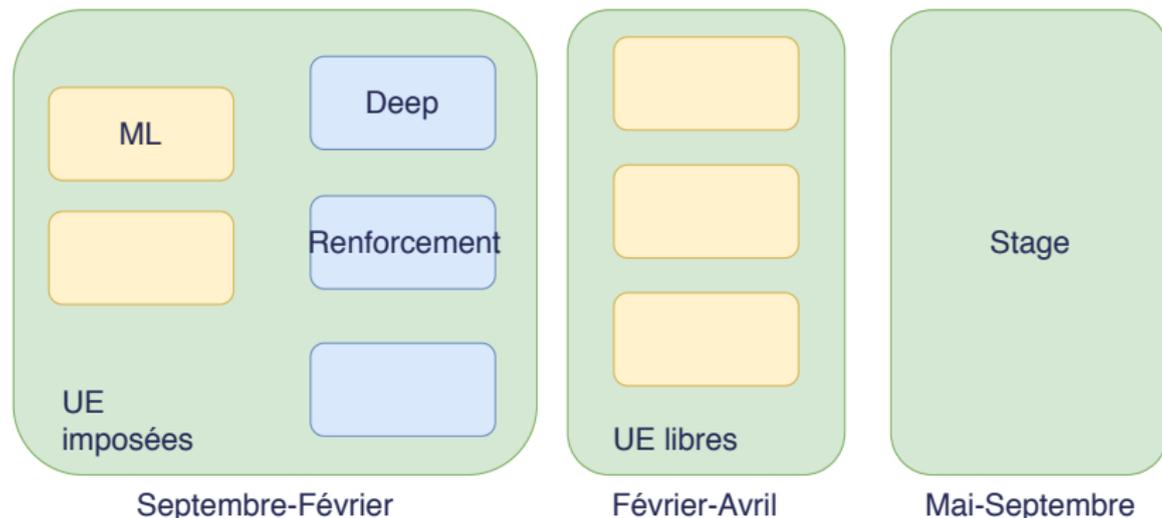
Gestion des
données

Analyse
statistique de
graphes

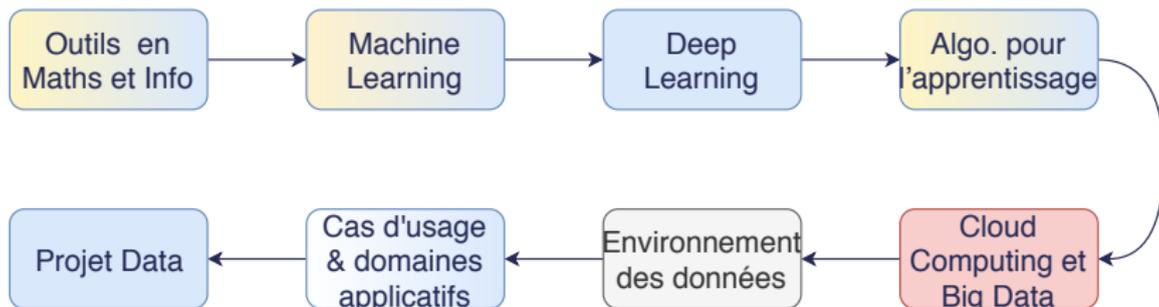
Algo. Stoch. : de
la finance aux
données
massives

<https://www.ljll.math.upmc.fr/FilBigData/>

Construction en cours d'un master à cheval entre les mathématiques et l'informatique. Début septembre 2019.



- UE du master de statistiques, UE du master DAC



- 2 jours d'encadrement par module (98h)
- Equipe stats, équipe MLIA, équipe BD, intervenants externes
- Public hétérogène
⇒ mise à niveau en programmation et en maths...
...mais un public très motivé

<https://formationmachinelearning.lip6.fr>

Conclusion

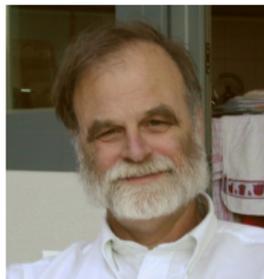
Sans informatique, pas de traitement de données

⇒ pas d'IA

Sans mathématiques, pas de création algorithmique

⇒ pas d'IA

- L'IA existe depuis aussi longtemps que l'informatique
- L'IA a déjà connu beaucoup de hauts et de bas
 - Les algorithmes/implémentations ont toujours progressé
 - Les avancées récentes sont significatives
 - Il est très difficile de prédire où on en sera dans 5 ans
- Crise des vocations académiques = problème majeur à court terme



Jim Gray – Prix Turing

“ Le traitement de données et la gestion de connaissances représentent ainsi le quatrième pilier de la science après la théorie, l'expérimentation et la simulation.”